Dis 5 Convergence of random variables

Wednesday, February 17, 2021

$X_1, X_2, X_3, \dots$ Q: 'get close' to some function?

$a_1, a_2, a_3 \dots \in \mathbb{R}$

1) almost sure $\Leftarrow$ SLLN

2) conv. in prob. $\Leftarrow$ WLLN

3) conv. in distri $\Leftarrow$ CLT

CLT: iid rv $X_1, X_2, \dots, X_n$, $E[X_i] = \mu$, $Var(X_i) = \sigma^2$

$$S_n \sim N(E[S_n], Var(S_n)), \text{ where } S_n = \sum_{i=1}^{n} X_i$$

$M_n = \frac{S_n}{n}$

$\Downarrow$

$$Z \sim \frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} \sim N(0,1)$$

$$P(Z \le z) = \Phi(z)$$

1. Confidence Interval Comparisons

In order to estimate the probability of a head in a coin flip, p, you flip a coin n times, where n is a positive integer, and count the number of heads, S_n. You use the estimator $\hat{p} = S_n/n$.

(a) You choose the sample size n to have a guarantee

$$\Pr(|\hat{p} - p| \ge \epsilon) \le \delta.$$

Using Chebyshev Inequality, determine n with the following parameters. Note that you should not have p in your final answer.

~~(i) Compare the value of n when $\epsilon = 0.05$, $\delta = 0.1$ to the value of n when $\epsilon = 0.1, \delta = 0.1$.~~

~~(ii) Compare the value of n when $\epsilon = 0.1, \delta = 0.05$ to the value of n when $\epsilon = 0.1, \delta = 0.1$~~

(b) Now, we change the scenario slightly. You know that $p \in (0.4, 0.6)$ and would now like to determine the smallest n such that

$$\Pr\left(\frac{|\hat{p} - p|}{p} \le 0.05\right) \ge 0.95.$$

Use the CLT to find the value of n that you should use. *Recall that the CLT states that the sum of IID random variables tends to a normal distribution with the sample mean and variance as it's parameters for n large enough.*

$$\frac{d}{dp} p(1-p) = 0$$

$$1 - 2p = 0 \Rightarrow p = \frac{1}{2}$$

a) $\Pr(|\underset{\hat{p}}{X} - \underset{p}{E(X)}| \ge \epsilon) \le \frac{Var(X)}{\epsilon^2}$

$$\frac{p(1-p)}{n\epsilon^2} \le \delta \Rightarrow \frac{1}{4\epsilon^2\delta} \le n$$

$E[\hat{p}] = \frac{1}{n} E[S_n] = \frac{1}{n} \times n \times p = p$

$S_i \sim Bern(p)$

$\Rightarrow \Pr(|\hat{p} - p| \ge \epsilon) \le \frac{Var(\hat{p})}{\epsilon^2}$

$Var(S_i) = \sigma^2 = p(1-p)$

$Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n) = \frac{1}{n^2} \times n\, Var(S_i) = \frac{1}{n}\sigma^2 \quad p(1-p)$

ε^2 $\quad Var(S_1) = \sigma^2 = p(1-p)$

$$Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n) = \frac{1}{n^2} \times n\, Var(S_1) = \frac{1}{n}\sigma^2 = \frac{p(1-p)}{n}$$

$n \to \infty$

$X_1, X_2, X_3, \ldots\ldots$

Convergence in probability

$X_n \xrightarrow{P} X$ ... "conv. in prob. to X"

$$\forall \varepsilon > 0, \quad \boxed{\lim_{n\to\infty} Pr(|X_n - X| > \varepsilon) = 0}$$

$$f(n) = \frac{1}{2^n}$$

2. Convergence of Exponentials

Let $X_1, X_2, \ldots$ be i.i.d. Exponential(λ) random variables. Show that

$$\frac{X_n}{\ln n} \to 0 \quad \text{in probability as } n \to \infty.$$

$f(n) \to 0$

$\lim_{n\to\infty}$

$n=1$ $\quad \ln 1$

$X_1' = \frac{X_1}{\ln 1}$

$X_2' = \frac{X_2}{\ln 2}$

$X_1' = \frac{X_1}{\ln 1}$

$X_2' = \frac{X_1}{\ln 2}$

$X_3' = \frac{X_1}{\ln 3}$ ..

Y_n $\quad \varepsilon \quad \varepsilon \quad 0$

Our sequence of RVs is $Y_1, Y_2, Y_3, \ldots$, where $Y_i = \frac{X_i}{\ln i}$

$$Y_n \xrightarrow{P} 0$$

$$\forall \varepsilon > 0, \lim_{n\to\infty} Pr(|Y_n - 0| > \varepsilon) = 0$$

$$\forall \varepsilon > 0, \quad Pr\left(\left|\frac{X_n}{\ln n} - 0\right| > \varepsilon\right) = Pr\left(\frac{X_n}{\ln n} > \varepsilon\right) = Pr(X_n > \varepsilon \ln n)$$

$$= 1 - CDF_{exp}(\varepsilon \ln n) = e^{-\lambda(\varepsilon \ln n)}$$

Almost sure convergence

$X_n \xrightarrow{a.s} X$

$\lim \quad e^{-\lambda(\varepsilon \ln n)}$

Almost sure convergence

$X_n \xrightarrow{as} X$

$\lim_{n\to\infty} e^{-\lambda(\varepsilon \ln n)} \to 0$

$F_X(x) = 1 - e^{-\lambda x}$

$P(\lim_{n\to\infty} X_n = X) = 1$

$\lim_{n\to\infty} a_n = a, \quad a_i \in R$

$\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$

$Pr((X_n = k), (X = k))$

$Pr(X_n = k) = Pr(X = k)$

$X_i \sim Bern(\frac{1}{2}) \to T, \to H$

$$X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X);$$
$$X_n \xrightarrow{p} X \Rightarrow g(X_n) \xrightarrow{p} g(X);$$
$$X_n \xrightarrow{a.s.} X \Rightarrow g(X_n) \xrightarrow{a.s.} g(X).$$

(continuous mapping theorem)

SLLN

$X_1, X_2, \ldots, X_n$ i.i.d. $E[X] = \mu$

$\frac{1}{n}\sum_{i=1}^{n} X_i \to E[X] = \mu$

$X_1, \ldots, X_n, \ldots \xrightarrow{p} X$

$X_1^2, X_2^2, \ldots \xrightarrow{p} X^2$

3. **Breaking a Stick**

I break a stick n times, where n is a positive integer, in the following manner: the ith time I break the stick, I keep a fraction X_i of the remaining stick where X_i is uniform on the interval $[0, 1]$ and $X_1, X_2, \ldots, X_n$ are i.i.d. Let $P_n = \prod_{i=1}^{n} X_i$ be the fraction of the original stick that I end up with.

(a) Show that $P_n^{1/n}$ converges almost surely to some constant function.

(b) Compute $\mathbb{E}[P_n]^{1/n}$.

$\lim_{n\to\infty} P_n^{\frac{1}{n}} = \square$

$\lim_{n\to\infty} \ln(P_n^{\frac{1}{n}}) = \lim_{n\to\infty} \frac{1}{n} \ln P_n = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \ln X_i$

$\xrightarrow{as} E[\ln X_1]$ by SLLN

$P_n^{\frac{1}{n}} \xrightarrow{a.s} \square \Rightarrow \lim_{n\to\infty} e^{\ln P_n^{\frac{1}{n}}} \xrightarrow{a.s} e^{E[\ln X_1]} = e^{-1}$

$(g(x) = e^x)$

$X_i \sim U(0,1)$

$E[\ln X_1] = \int_{x=0}^{x=1} \cancel{f_X(x)} \ln x \, dx$

$= \int_0^1 \ln x \, dx$

$= [x \ln x - x]_0^1$

$= 0 - 1 = -1$

Markov's: $X > 0$, $Pr(X \geq c) \leq E[X]/c$

Chebyshev's: $Pr(|X - E[X]| \geq c) \leq Var[X]/c^2$

WLLN: $X_1, X_2, \ldots, X_n$ w/ $E[X_i] = \mu$, $Var[X_i] = \sigma^2$ IID

define $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $E[M_n] = \mu$, $Var[M_n] = \frac{\sigma^2}{n}$

$Pr(|M_n - E[M_n]| \geq c) \leq \frac{\sigma^2}{n c^2}$

n define $M_n = \frac{1}{n} \sum_{i=1} X_i$, $E[M_n] = \mu$, $var[M_n] = \frac{}{n}$

$$\Pr(|M_n - E[M_n]| \geq c) \leq \frac{\sigma^2}{nc^2}$$

empirical mean   true mean

$$\lim_{n \to \infty} \Pr(|M_n - \mu| \geq c) \leq \frac{\sigma^2}{nc^2} \xrightarrow{n \to \infty} 0$$

$= 6 - 7 = -1$